

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:21:03 ; Search time 1738 Seconds
(without alignments)
1784.312 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64

Sequence: 1 tcagacccatttaggac.....ttgttagcttttctctaga 64

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	245095	2	AC111774
2	61.4	95.9	4702	10	BC060227
3	61.4	95.9	4940	10	BC065115
C 4	61.4	95.9	141790	10	AL671759
5	45.6	71.2	1477	9	AF070546
C 6	45.6	71.2	2710	11	BV178942
7	45.6	71.2	2944	9	BC065263
8	45.6	71.2	3049	9	AK096161
9	45.6	71.2	3495	9	BC029686
10	45.6	71.2	3554	9	BC048565
11	45.6	71.2	4809	9	BC080578
12	45.6	71.2	4880	9	HSM803946
13	45.6	71.2	5023	9	HSM803947
14	45.6	71.2	22255	6	AR409342
15	45.6	71.2	22255	6	AX239608
16	45.6	71.2	115755	9	HSDJ62215
17	41.4	64.7	4722	6	AR199533
18	41.4	64.7	4722	6	AR374685
19	41.4	64.7	4722	6	AR409320

20	41.4	64.7	4722	6	AX239571	AX239571 Sequence
21	41.4	64.7	4722	6	BD056447	BD056447 Novel low
22	39.2	61.3	4697	6	AR199536	AR199536 Sequence
23	39.2	61.3	4697	6	AR374688	AR374688 Sequence
24	39.2	61.3	4697	6	AR409323	AR409323 Sequence
25	39.2	61.3	4697	6	AX239574	AX239574 Sequence
26	39.2	61.3	4697	6	BD056450	BD056450 Novel low
27	31	48.4	204697	2	AC120995	AC120995 Rattus no
C 28	31	48.4	236070	2	AC112571	AC112571 Rattus no
C 29	30	46.9	155580	2	AC118845	AC118845 Rattus no
30	29.6	46.2	229213	10	AL772138	AL772138 Mouse DNA
31	29.2	45.6	218924	10	AC124567	AC124567 Mus muscu
C 32	29	45.3	133265	10	AL954355	AL954355 Mouse DNA
C 33	29	45.3	142687	10	AL135758	AL135758 Mouse DNA
C 34	29	45.3	216851	2	AC118154	AC118154 Rattus no
35	29	45.3	220407	2	AC121626	AC121626 Rattus no
36	29	45.3	233944	2	AC106583	AC106583 Rattus no
C 37	28.8	45.0	3742	5	BC077481	BC077481 Xenopus 1
C 38	28.6	44.7	170788	2	AC139387	AC139387 Rattus no
C 39	28.6	44.7	244536	2	AC099176	AC099176 Rattus no
C 40	28.6	44.7	334320	2	AC113256	AC113256 Rattus no
C 41	28.4	44.4	378	6	CQ656202	CQ656202 Sequence
42	28.4	44.4	182389	9	AP002829	AP002829 Homo sapi
43	28.4	44.4	183090	2	AC121175	AC121175 Rattus no
C 44	28.4	44.4	196832	9	AC018694	AC018694 Homo sapi
C 45	28.4	44.4	208326	2	AC022888	AC022888 Homo sapi

ALIGNMENTS

RESULT 1
AC111774
LOCUS
DEFINITION
AC111774
Rattus norvegicus clone CH230-171G2, WORKING DRAFT SEQUENCE, 10
unordered pieces.
AC111774
VERSION
AC111774.4 GI:24818709
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 245095)
Muzny,D,Marle., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Chen,Y., Chen,Z., Chu,J.,
Chacko,J., Chavaz,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto-M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gregeorge,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivat,A.,
Karpach,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulsegged,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martineez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajda, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wleczky, R., Wooden, H., Worley, K., Williams, G., Willson, R., Wleczky, R., Wleczky, H., Wleczky, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 245095)
Morley, K.C.

Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 245095)

Rat Genome Sequencing Consortium.
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2002 this sequence version replaced gi:23321701.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOBQ
Center clone name: CH230-171G2
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 153773 bases at least Q40
Consensus quality: 160170 bases at least Q30
Consensus quality: 164659 bases at least Q20
Estimated insert size: 160182; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
69762: contig of 69762 bp in length
69862: gap of unknown length
69863: contig of 3827 bp in length
73789: gap of unknown length
73790: contig of 14658 bp in length
88447: gap of unknown length
88448: contig of 36066 bp in length
124613: contig of 36066 bp in length
124713: gap of unknown length
179270: contig of 54557 bp in length
179271: gap of unknown length
179370: contig of 45126 bp in length
224496: gap of unknown length
224597: contig of 3540 bp in length
228137: gap of unknown length
228237: contig of 11229 bp in length
239465: gap of unknown length
239566: contig of 4277 bp in length
243842: gap of unknown length
243943: contig of 1153 bp in length.
243943: Location/Qualifiers
1..245095
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-171G2"
88548..89924
/note="wgs_contig"
91432..93819
/note="wgs_contig"
116848..118080
/note="wgs_contig"
124714..126005
/note="wgs_contig"
228237..229783
/note="wgs_contig"
239566..240945
/note="wgs_contig"

FEATURES
Source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
ORIGIN
Query Match 100.0%; Score 64; DB 2; Length 245095;
Best Local Similarity 100.0%; Pred. No. 6.3e-11;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCATGACCTCATTTAGGACCAAGAGCTGTGGTCTTAGATCTTCTTC 60
Db 148400 TCATGACCTCATTTAGGACCAAGAGCTGTGGTCTTAGATCTTCTTC 148459
Qy 61 TAGA 64
Db 148460 TAGA 148463
RESULT 2
BC060227 Mus musculus taxilin, mRNA (cdna clone MGC:77972 IMAGE:30357541), complete cds.
ACCESSION BC060227
VERSION BC060227.1 GI:38614270
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
AUTHORS
1 (bases 1 to 4702)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:20:29; Search time 266 Seconds
(without alignments)
1424.299 Million cell updates/sec

Title: US-10-663-418-71
Perfect score: 64
Sequence: 1 tcatgacatcttttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N Geneseq 16Dec04:
1: Geneseqn1980s:
2: Geneseqn1990s:
3: Geneseqn2000s:
4: Geneseqn2001as:
5: Geneseqn2001bs:
6: Geneseqn2002as:
7: Geneseqn2002bs:
8: Geneseqn2003as:
9: Geneseqn2003bs:
10: Geneseqn2003cs:
11: Geneseqn2003ds:
12: Geneseqn2004as:
13: Geneseqn2004bs:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.6	71.2	1926	6 ABK35029	Abk35029 Human cDN
2	45.6	71.2	2145	13 ADP54990	Adp54990 Human PRO
3	45.6	71.2	5085	8 ACC46153	Acc46153 Human dit
4	45.6	71.2	6577	8 ACC46130	Acc46130 Human dit
5	45.6	71.2	22255	5 AAH26497	Aah26497 Human low
6	45.6	71.2	115756	8 ACD13448	Acd13448 Human DNA
7	41.4	64.7	4722	2 AAH26491	Aah26491 Rabbit lo
8	41.4	64.7	4722	5 AAH26491	Aah26491 Rabbit lo
9	39.2	61.3	4697	2 AAH26496	Aah26496 Human low
10	39.2	61.3	4697	5 AAH26496	Aah26496 Human low
11	38	59.4	635	13 ADQ52076	Adq52076 Novel can
12	33.6	52.5	268	2 AAT23888	Aat23888 Human gen
13	27.2	42.5	1024	6 ABX66285	Abx66285 Helicobac
14	27.2	42.5	1024	6 ABX65642	Abx65642 Helicobac
15	27.2	42.5	1024	6 ABX66280	Abx66280 Helicobac
16	27.2	42.5	1035	2 AAX14487	Aax14487 H. pylori
17	27.2	42.5	6325	8 ABZ10086	Abz10086 Haematopo
18	26.4	41.2	569	11 ACN87401	Acn87401 Breast ca
19	26.4	41.2	2418	6 ABQ70768	Abq70768 Listeria
20	26.4	41.2	2875	12 ADM91308	Adm91308 DNA homol

C 21	26.4	41.2	25800	12	ADQ97464	Adq97464 Mouse can
C 22	26.4	41.2	110000	6	ABA03041_21	Continuation (22 o
C 23	26.2	40.9	2730	2	AAT32583	Aat32583 Alfalfa i
24	26.2	40.9	15690	6	ABK15038	Abk15038 Canine di
25	26.2	40.9	15690	6	ABK15000	Abk15000 Canine di
26	26.2	40.9	18826	6	ABK15039	Abk15039 Canine di
27	25.8	40.3	16825	4	ABK42608	Abk42608 Genomic s
28	25.8	40.3	16825	9	ADB60764	Adb60764 Connectiv
29	25.8	40.3	23243	9	ADA02567	Ada02567 Mouse Gat
30	25.8	40.3	23243	10	ADB72305	Adb72305 Mouse Gat
31	25.8	40.3	23243	10	ADE95815	Ade95815 Mouse Gat
32	25.8	40.3	26997	4	AAS46748	Aas46748 Tumour su
33	25.8	40.3	192427	10	ADL13825	Adl13825 Osteoarth
34	25.8	40.3	310268	13	ABD32548	Abd32548 Human can
35	25.6	40.0	1754	10	ADB63505	Adb63505 Human cDN
36	25.6	40.0	6325	8	ABZ10232	Abz10232 Haematopo
37	25.6	40.0	110000	4	AAK95240_05	Continuation (6 of
38	25.6	40.0	110000	4	AAK96733_05	Continuation (6 of
39	25.6	40.0	110000	6	ABT00010_05	Continuation (6 of
40	25.6	40.0	110000	6	ABT01503_05	Continuation (6 of
41	25.6	40.0	110000	12	ADH77486_05	Continuation (6 of
42	25.4	39.7	28449	12	ADQ97240	Adq97240 Mouse can
43	25.4	39.7	256294	13	ABD33020	Abd33020 Mouse can
44	25.2	39.4	275	4	AAK83205	Aak83205 Human imm
45	25.2	39.4	427	6	ABL38563	Abi38563 Human col

ALIGNMENTS

RESULT 1

ABK35029
ID ABK35029 standard; cDNA; 1926 BP.
XX
AC ABK35029;
DT 08-MAY-2002 (first entry)
XX
DE Human cDNA encoding secreted protein #167.

Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
viral infection; bacterial infection; fungal infection; diabetes; asthma;
autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
coagulation disorder; inflammatory disorder; Crohn's disease; incision;
tissue regeneration; wound healing; burn; haematopoiesis;
myeloid cell deficiency; lymphoid cell deficiency.

OS Homo sapiens.

XX WO200177288-A2.

PD 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US010224.

XX 06-APR-2000; 2000US-0195582P.

XX (GEMV) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;

XX Gulukota K, Graham JR;

XX WPI; 2002-179321/23.

XX Five hundred and ninety two polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for treating
XX immune deficiencies and disorders such as autoimmune disorders.

XX Claim 1; Page 152; 372pp; English.

XX The invention relates to 592 polynucleotides which have been derived from

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:25:03 ; Search time 97 Seconds
(without alignments)
1079.605 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64

Sequence: 1 tcatgacctcatttaggac.....ttgttagcttttctctaga 64

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:**
1: /cgm2_6/ptodata/1/ina/5A COMB.seq:**
2: /cgm2_6/ptodata/1/ina/5B COMB.seq:**
3: /cgm2_6/ptodata/1/ina/6A COMB.seq:**
4: /cgm2_6/ptodata/1/ina/6B COMB.seq:**
5: /cgm2_6/ptodata/1/ina/PCTUS COMB.seq:**
6: /cgm2_6/ptodata/1/ina/backfiles.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.6	71.2	22255	4	US-09-616-289-51
2	41.4	64.7	4722	3	US-08-979-608A-14
3	41.4	64.7	4722	4	US-09-517-849-14
4	41.4	64.7	4722	4	US-09-616-289-14
5	39.2	61.3	4697	3	US-08-979-608A-17
6	39.2	61.3	4697	4	US-09-517-849-17
7	39.2	61.3	4697	4	US-09-616-289-17
8	28.8	45.0	601	4	US-09-949-016-171883
9	27.6	43.1	601	4	US-09-949-016-171884
10	27.6	43.1	232547	4	US-09-949-016-16603
11	26.4	41.2	601	4	US-09-949-016-149477
12	26.4	41.2	156942	4	US-09-949-016-12227
13	26.4	41.2	156950	4	US-09-949-016-15946
14	26.4	41.2	300598	4	US-09-949-016-11868
15	26.4	41.2	302604	4	US-09-949-016-14588
16	26.4	41.2	302604	4	US-09-949-016-14589
17	26.4	41.2	308362	4	US-09-949-016-17119
18	26.2	40.9	2730	1	US-08-339-129-1
19	25.8	40.3	601	4	US-09-949-016-158019
20	25.8	40.3	601	4	US-09-949-016-158020
21	25.8	40.3	36620	4	US-09-949-016-16150
22	25.8	40.3	87352	4	US-09-949-016-12053
23	25.8	40.3	87352	4	US-09-949-016-12721
24	25.8	40.3	87352	4	US-09-949-016-15692
25	25.8	40.3	87352	4	US-09-949-016-15693
26	25.4	39.7	601	4	US-09-949-016-124839
27	25.4	39.7	601	4	US-09-949-016-124840

Sequence 15777, A
Sequence 15270, A
Sequence 12583, A
Sequence 17392, A
Sequence 11, Appl
Sequence 130194, A
Sequence 15419, A
Sequence 24772, A
Sequence 24773, A
Sequence 24774, A
Sequence 35142, A
Sequence 35143, A
Sequence 35144, A
Sequence 35171, A
Sequence 35172, A
Sequence 35173, A
Sequence 173472, A
Sequence 173473, A

28 25.4 39.7 80411 4 US-09-949-016-15777
29 25.4 39.7 192302 4 US-09-949-016-15270
c 30 25.2 39.4 254964 4 US-09-949-016-12583
c 31 25.2 39.4 254964 4 US-09-949-016-17392
32 25.2 39.4 392000 4 US-10-027-983-11
33 25 39.1 601 4 US-09-949-016-130194
34 25 39.1 212449 4 US-09-949-016-15419
c 35 24.8 38.8 601 4 US-09-949-016-24772
c 36 24.8 38.8 601 4 US-09-949-016-24773
c 37 24.8 38.8 601 4 US-09-949-016-24774
c 38 24.8 38.8 601 4 US-09-949-016-35142
c 39 24.8 38.8 601 4 US-09-949-016-35143
c 40 24.8 38.8 601 4 US-09-949-016-35144
c 41 24.8 38.8 601 4 US-09-949-016-35171
c 42 24.8 38.8 601 4 US-09-949-016-35172
c 43 24.8 38.8 601 4 US-09-949-016-35173
c 44 24.8 38.8 601 4 US-09-949-016-173472
c 45 24.8 38.8 601 4 US-09-949-016-173473

ALIGNMENTS

RESULT 1

US-09-616-289-51
; Sequence 51, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-51

Query Match 71.2%; Score 45.6; DB 4; Length 22255;
Best Local Similarity 92.3%; Pred. No. 7.6e+06;
Matches 48; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 12 TTTTAGACCAAGAGCTGCTGCTTTCTTAGATTGTTAGCTTTTCTCTAG 63

Db 19382 TTTTAGACCAAGATCTGCTGCTTTCTTAGATTGTTAGCTTTTCTCTAG 19433

RESULT 2

US-08-979-608A-14
; Sequence 14, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:58:43 ; Search time 300 Seconds
(without alignments)
1295.158 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64
Sequence: 1 tcatgacctatttaggac.....ttgttagcttttctctaga 64

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq:*

13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*

20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

22: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

23: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	45.6	71.2	1926	US-09-822-849A-167
2	45.6	71.2	22255	US-09-976-740-51
3	45.6	71.2	22255	US-10-023-529-51
4	45.6	71.2	22255	US-10-023-523-51
5	45.6	71.2	22255	US-10-616-187-51
6	45.6	71.2	22255	US-10-671-242-51
7	41.4	64.7	4722	US-09-962-055-14
8	41.4	64.7	4722	US-09-976-740-14
9	41.4	64.7	4722	US-10-023-529-14
10	41.4	64.7	4722	US-10-023-523-14
11	41.4	64.7	4722	US-10-616-187-14

12	41.4	64.7	4722	17	US-10-671-242-14	Sequence 14, Appl
13	39.2	61.3	4697	9	US-09-962-055-17	Sequence 17, Appl
14	39.2	61.3	4697	9	US-09-976-740-17	Sequence 17, Appl
15	39.2	61.3	4697	13	US-10-023-529-17	Sequence 17, Appl
16	39.2	61.3	4697	13	US-10-023-523-17	Sequence 17, Appl
17	39.2	61.3	4697	17	US-10-616-187-17	Sequence 17, Appl
18	39.2	61.3	4697	17	US-10-671-242-17	Sequence 17, Appl
c 19	28.4	44.4	378	17	US-10-242-535A-1128	Sequence 1128, Ap
c 20	28.4	44.4	378	17	US-10-085-783A-1128	Sequence 1128, Ap
c 21	27.2	42.5	1035	10	US-09-882-227-401	Sequence 401, App
c 22	27.2	42.5	6325	18	US-10-473-126-226	Sequence 226, App
c 23	27.2	42.5	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 24	26.8	41.9	818	13	US-10-027-632-152420	Sequence 152420,
c 25	26.8	41.9	818	17	US-10-027-632-152420	Sequence 152420,
c 26	26.4	41.2	569	14	US-10-198-846-8551	Sequence 8551, Ap
c 27	26.4	41.2	739	13	US-10-027-632-171545	Sequence 171545,
c 28	26.4	41.2	739	17	US-10-027-632-171545	Sequence 171545,
c 29	26.4	41.2	2418	17	US-10-398-221-3581	Sequence 3581, Ap
c 30	26.2	40.9	673	13	US-10-027-632-7733	Sequence 7733, Ap
c 31	26.2	40.9	673	13	US-10-027-632-7734	Sequence 7734, Ap
c 32	26.2	40.9	673	13	US-10-027-632-7735	Sequence 7735, Ap
c 33	26.2	40.9	673	17	US-10-027-632-7733	Sequence 7733, Ap
c 34	26.2	40.9	673	17	US-10-027-632-7734	Sequence 7734, Ap
c 35	26.2	40.9	673	17	US-10-027-632-7735	Sequence 7735, Ap
c 36	26	40.6	492	13	US-10-027-632-267996	Sequence 267996,
c 37	26	40.6	492	13	US-10-027-632-267997	Sequence 267997,
c 38	26	40.6	492	17	US-10-027-632-267996	Sequence 267996,
c 39	26	40.6	492	17	US-10-027-632-267997	Sequence 267997,
c 40	25.8	40.3	16825	9	US-09-764-847-1495	Sequence 1495, Ap
c 41	25.8	40.3	16825	14	US-10-092-154-1495	Sequence 1495, Ap
c 42	25.8	40.3	23243	17	US-10-052-482-73	Sequence 73, Appl
c 43	25.8	40.3	26997	17	US-10-221-714A-474	Sequence 474, App
c 44	25.8	40.3	301692	17	US-10-428-487-11	Sequence 11, Appl
c 45	25.8	40.3	310268	18	US-10-367-094-195	Sequence 195, App

ALIGNMENTS

RESULT 1

US-09-822-849A-167

Sequence 167, Application US/09822849A

Patent No. US20020045170A1

GENERAL INFORMATION:

APPLICANT: Wong, Gordon G.

APPLICANT: Fechtel, Kim

APPLICANT: Agostino, Michael J.

APPLICANT: Howes, Steven H.

APPLICANT: Resnick, Richard J.

APPLICANT: Gulukota, Kamalak

APPLICANT: Graham, James R.

APPLICANT: Genetics Institute, Inc.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS

FILE REFERENCE: GIN 6403

CURRENT APPLICATION NUMBER: US/09/822,849A

CURRENT FILING DATE: 2001-09-04

PRIOR APPLICATION NUMBER: 60/195,582

PRIOR FILING DATE: 2000-04-06

NUMBER OF SEQ ID NOS: 598

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 167

LENGTH: 1926

TYPE: DNA

ORGANISM: Homo sapiens

US-09-822-849A-167

Query Match 71.2%; Score 45.6; DB 9; Length 1926;

Best Local Similarity 92.3%; Pred. No. 1e-05; Indels 0; Gaps 0;

Matches 48; Conservative 0; Mismatches 4;

12 TTTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTTCTCTAG 63

|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:21:49 ; Search time 1820 Seconds
(without alignments)
1338.524 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64

Sequence: 1 tcagacacatttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_g881:*

9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	64	100.0	566	AA925303	AA925303 UI-R-A1-e
C 2	64	100.0	566	BQ192706	BQ192706 UI-R-DL1-
C 3	64	100.0	566	CA503893	CA503893 UI-R-FU0-
C 4	64	100.0	566	CB322784	CB322784 UI-R-DY0-
C 5	64	100.0	802	CO395194	CO395194 AGENCOURT
C 6	63	98.4	419	BF555411	BF555411 UI-R-A1-e
C 7	63	98.4	477	AA869795	AA869795 vq15h08.1
C 8	63	98.4	546	EX516797	EX516797 BX516797
C 9	63	98.4	641	CA880144	CA880144 K0980808-
C 10	63	98.4	641	CB058161	CB058161 NISC_jx01
C 11	63	98.4	1009	CB208950	CB208950 AGENCOURT
C 12	63	98.4	1167	BF180039	BF180039 601806391
C 13	61.4	95.9	519	AA170252	AA170252 ms87e12.1
C 14	61.4	95.9	540	AA915373	AA915373 v229d01.1
C 15	61.4	95.9	647	CA450819	CA450819 UI-M-FY0-
C 16	61.4	95.9	651	BP762179	BP762179 BP762179
C 17	61.4	95.9	651	CF727202	CF727202 UI-M-HB0-
C 18	61.4	95.9	652	BQ176939	BQ176939 UI-M-DJ2-
C 19	61.4	95.9	655	BP761959	BP761959 BP761959
C 20	61.4	95.9	675	CO042840	CO042840 UI-M-GH0-
C 21	61.4	95.9	694	BF120794	BF120794 601758189
C 22	61.4	95.9	698	AK011182	AK011182 Mus muscu
C 23	61.4	95.9	712	BY710509	BY710509 BY710509
C 24	61.4	95.9	717	CA324014	CA324014 UI-M-FY0-

C 25	61.4	95.9	748	7	CO042637	CO042637 UI-M-GH0-
C 26	61.4	95.9	760	7	CO042337	CO042337 UI-M-GH0-
C 27	61.4	95.9	856	2	BB038251	BB038251 BB038251
C 28	59.8	93.4	454	6	BY597346	BY597346 BY597346
C 29	59.8	93.4	543	5	BU609826	BU609826 UI-M-DJ2-
C 30	59.8	93.4	755	7	CN529567	CN529567 UI-M-HQ0-
C 31	54.4	85.0	174	1	AI608525	AI608525 VC73b12.1
C 32	53.4	83.4	650	4	BG867325	BG867325 602788676
C 33	47.2	73.8	309	2	AW487474	AW487474 84843 MAR
C 34	47.2	73.8	340	2	AW446148	AW446148 84459 MAR
C 35	47.2	73.8	481	2	BE755989	BE755989 209991 MA
C 36	47.2	73.8	482	4	BM090178	BM090178 505329 MA
C 37	47.2	73.8	513	4	BI774054	BI774054 466018 MA
C 38	47.2	73.8	640	7	CK953122	CK953122 4092825 B
C 39	47.2	73.8	667	7	CK950080	CK950080 4075226 B
C 40	47.2	73.8	692	7	CK953458	CK953458 4093209 B
C 41	46.2	72.2	548	1	AV604019	AV604019 AV604019
C 42	45.6	71.2	364	7	T99127	T99127 ve62d08.r1
C 43	45.6	71.2	432	4	BM766498	BM766498 K-EST0048
C 44	45.6	71.2	489	4	BM689900	BM689900 UI-E-CK1-
C 45	45.6	71.2	505	2	BE185069	BE185069 MRI-HT070

ALIGNMENTS

RESULT 1
AA925303/c
LOCUS
DEFINITION
AA925303 566 bp mRNA linear EST 04-JUL-1999
UI-R-A1-ek-e-09-0-UI.s1 UI-R-A1 Rattus norvegicus cDNA clone
UI-R-A1-ek-e-09-0-UI 3' similar to gi|1768059|gb|AA184410|AA184410
mt34f05.r1 Soares mouse 3NBMS Mus musculus cDNA clone 622977 5',
mRNA sequence.
ACCESSION
AA925303
VERSION
AA925303.1 GI:4236494
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 566)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
97044477
8889548
COMMENT
On Apr 21, 1998 this sequence version replaced gi:3072439.
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence tag present in the cDNA between the NotI site and the
oligo-tr track served to identify it as a clone from the normalized
adult muscle library. cDNA Library Preparation: M. Fatima Bonaldo,
Ph.D. Clone distribution: clones will be available through Research
Genetics This clone is also available through the I.M.A.G.E.
Consortium at LNL (info@image.llnl.gov). IMAGE ID=1771492
Seq primer: M13 Forward
POLYA=NO. Location/Qualifiers
1. 566
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-A1-ek-09-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"

/clone_lib="UI-R-A1"
 /note="vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-A1
 library is a subtracted library derived from the UI-R-A0
 library. The UI-R-A0 library consisted of a mixture of
 individually tagged normalized libraries constructed from
 rat placenta, adult lung, brain, liver, kidney, heart,
 spleen, ovary, and muscle. The tag is a string of 3-5
 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (UI-R-A1) was constructed as follows: PCR
 amplified cDNA inserts from a pool of approximately 3,840
 UI-R-A0 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-A0
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-A1 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

ORIGIN

Query Match 100.0%; Score 64; DB 1; Length 566;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGACCTCATTTTAGGACCAAGAGCTGTGTTCTTTAGATTGTTAGCTTTTCTC 60
 Db 565 TCATGACCTCATTTTAGGACCAAGAGCTGTGTTCTTTAGATTGTTAGCTTTTCTC 506
 Qy 61 TAGA 64
 Db 505 TAGA 502

RESULT 2

BQ192706/c
 LOCUS BQ192706 649 bp mRNA linear EST 30-APR-2002
 DEFINITION UI-R-DRI-cla-d-11-0-UI-s1 UI-R-DRI Rattus norvegicus cDNA clone
 UI-R-DRI-cla-d-11-0-UI 3', mRNA sequence.

ACCESSION BQ192706

VERSION BQ192706.1 GI:20368257

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

1 (bases 1 to 649)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL 97044477

MEDLINE 8889548

PUBMED

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: bento-soares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 normalized osteoblast library cDNA Library Preparation: M.B. Soares
 Lab Clone distribution: clones will be available through Research
 Genetics (www.reagen.com)

Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

source Location/Qualifiers
 1..649
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-DRI-cla-d-11-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-DRI"
 /note="vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-DRI
 library is a normalized Rat Osteoblast library (nRBO)
 constructed in pT737 vector according to the procedure
 described by Bonaldo, Lennon & Soares (Normalization and
 Subtraction: Two Approaches to Facilitate Gene Discovery.
 Genome Research 6: 791-806, 1996). The oligonucleotide
 used to prime first strand synthesis contained the
 sequence tag AAGATATCAA between the Not I cloning site and
 dT18 stretch. The Rat Osteoblast tissue was provided by
 Lian & Stein of the University of Massachusetts Medical
 School.
 TAG TISSUE=osteoblast
 TAG_LIB=UI-R-DRI
 TAG_SEQ=AAGATATCAA"

ORIGIN

Query Match 100.0%; Score 64; DB 5; Length 649;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCATGACCTCATTTTAGGACCAAGAGCTGTGTTCTTTAGATTGTTAGCTTTTCTC 60
 Db 575 TCATGACCTCATTTTAGGACCAAGAGCTGTGTTCTTTAGATTGTTAGCTTTTCTC 516
 Qy 61 TAGA 64
 Db 515 TAGA 512

RESULT 3

CA503893/c

LOCUS

DEFINITION

UI-R-FJ0-cpv-o-09-0-UI-s1 UI-R-FJ0 Rattus norvegicus cDNA clone

UI-R-FJ0-cpv-o-09-0-UI 3', mRNA sequence.

CA503893

CA503893.1 GI:24994847

KEYWORDS EST.

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 659)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL 97044477

MEDLINE 8889548

PUBMED

COMMENT

Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source
Location/Qualifiers
1. .659
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJO-cpv-0-09-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-FJO"
/note="Vector: pYX-ASC; Site 1: EcoR I; Site 2: Not I; UI-R-FJO is a cDNA library containing the following tissue(s): rat embryo. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pYX-ASC vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CATCTACT. This library was created for the University of Iowa Program for Rat Gene Discovery and Mapping (Val Sheffield, Bento Soares and Tom Casavant)
TAG_TISSUE=rat-embryo
TAG_LIB=UI-R-FJO
TAG_SEQ=CATCTACT

Query Match 100.0%; Score 64; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;
QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 60
DB 569 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 510
QY 61 TAGA 64
DB 509 TAGA 506

ORIGIN

Query Match 100.0%; Score 64; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;
QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 60
DB 569 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 510
QY 61 TAGA 64
DB 509 TAGA 506

RESULT 4
CB322784/c
LOCUS
DEFINITION
UI-R-DYO-crc-d-14-0-UI.e1 NCI CGAP_DYO Rattus norvegicus cDNA clone
IMAGE:7329856 3', mRNA sequence.
ACCESSION
CB322784
VERSION
CB322784.1 GI:28857442
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 683)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene discovery
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
9704477
8889548
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-re@mail.nih.gov

Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone Distribution: Clones will be available through IMAGE (http://image.llnl.gov)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
Source
Location/Qualifiers
1. .683
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="IMAGE:7329856"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_DYO"
/note="Vector: pYX-ASC; Site 1: EcoR I; Site 2: Eco RI; UI-R-DYO is a non-normalized Rat cartilage library (RC) constructed in pYX-ASC vector according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag CTAATGGACG between the Not I cloning site and dT18 stretch. The Rat cartilage tissue was provided by Dr Jeff Stevens at the University of Iowa.
TAG_TISSUE=cartilage
TAG_LIB=UI-R-DYO
TAG_SEQ=CTAATGGACG

Query Match 100.0%; Score 64; DB 6; Length 683;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;
QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 60
DB 570 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 511
QY 61 TAGA 64
DB 510 TAGA 507

ORIGIN

Query Match 100.0%; Score 64; DB 6; Length 683;
Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;
QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 60
DB 570 TCATGACCTCATTTAGGACCAAGAGCTGTGTGTTCTTAGATTGTTAGTTTTC 511
QY 61 TAGA 64
DB 510 TAGA 507

RESULT 5
CO395194
LOCUS
DEFINITION
AGENCOURT 27849307 NIH MGC 252 Rattus norvegicus cDNA clone
IMAGE:7310479 5', mRNA sequence.
ACCESSION
CO395194
VERSION
CO395194.1 GI:49577110
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 802)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
JOURNAL
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: LLM15356 row: m column: 05
 High quality sequence stop: 654.
 Location/Qualifiers
 1. .802
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clones="IMAGE:7310479"
 /tissue_type="Ovary, strain - Norway Line3. Age8 weeks.
 Tissue was snap-frozen adn transferred in -70C. RNase free
 for the entire procedure"
 /lab_hosts="DH10B Tona"
 /clone_lib="NIH_MGC_252"
 /notes="Organ: ovary; Vector: pDONR 201; Site 1: NotI;
 Site 2: NotI; RNA obtained from female ovaries animals at
 8 wk old. Tissues were snap-frozen and kept at -80C for
 two days before RNA extraction and purification
 (TRI-reagent method). cDNA was primed using oligo-dT
 primer: 5'-pGACTAGTCTAGTCGCGCGCGCCG(T)25-3' and
 cloned into the EcoRV/NotI sites of pExpress-1.
 Size-selection >1.25kb resulted in an average insert size
 of 1.7kb. This primary library is not normalized
 (normalized library is NIH_MGC 252) and was constructed by
 Express Genomics (Frederick, MD). Note: this is a NIH_MGC
 Library"

ORIGIN

Query Match 100.0%; Score 64; DB 7; Length 802;
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TCATGACCTCATTATAGACCAAGAGCTGTGTTCTTAGATGTTAGCTTTCTC 60
 Db 536 TCATGACCTCATTATAGACCAAGAGCTGTGTTCTTAGATGTTAGCTTTCTC 595
 Qy 61 TAGA 64
 Db 596 TAGA 599

RESULT 6
 BF555411
 LOCUS 419 bp mRNA linear EST 12-DEC-2000
 DEFINITION UI-R-A1-ek-e-09-0-UI.r1 UI-R-A1 Rattus norvegicus cDNA clone
 UI-R-A1-ek-e-09-0-UI 5', mRNA sequence.
 ACCESSION BF555411
 VERSION BF555411.1 GI:11665141
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 419)
 REFERENCE 1 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL 97044477
 MEDLINE 8889548
 PUBMED
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEERF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LLM (info@image.llnl.gov). IMAGE ID- 1771492
 Seq primer: M13 Forward.

FEATURES
 source

Location/Qualifiers
 1. .419
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-A1-ek-e-09-0-UI"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UI-R-A1"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-A1
 library is a subtracted library derived from the UI-R-A0
 library. The UI-R-A0 library consisted of a mixture of
 individually tagged normalized libraries constructed from
 rat placenta, adult lung, brain, liver, kidney, heart,
 spleen, ovary, and muscle. The tag is a string of 3-5
 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (UI-R-A1) was constructed as follows: PCR
 amplified cDNA inserts from a pool of approximately 3,840
 UI-R-A0 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-A0
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-A1 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

ORIGIN

Query Match 98.4%; Score 63; DB 2; Length 419;
 Best Local Similarity 98.4%; Pred. No. 4.2e-10;
 Matches 63; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 TCATGACCTCATTATAGACCAAGAGCTGTGTTCTTAGATGTTAGCTTTCTC 60
 Db 308 TCATGACCTCATTATAGACCAAGAGCTGTGTTCTTAGATGTTAGCTTTCTC 367
 Qy 61 TAGA 64
 Db 368 TAGA 371

RESULT 7

AA869795
 LOCUS 477 bp mRNA linear EST 16-MAR-1998
 DEFINITION vql5h08.r1 Barstead stromal cell line MFLRB8 Mus musculus cDNA
 clone IMAGE:1094367 5', mRNA sequence.
 ACCESSION AA869795
 VERSION AA869795.1 GI:2965240
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 477)
 REFERENCE 1 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:21:03 ; Search time 1738 Seconds
(without alignments)
1784.312 Million cell updates/sec

Title: US-10-663-418-71
Perfect score: 64
Sequence: 1 tcatgacctcattttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.hig.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	64	100.0	245095	2 AC111774	AC111774 Rattus no
2	61.4	95.9	4702	10 BC060227	BC060227 Mus muscu
3	61.4	95.9	4940	10 BC065115	BC065115 Mus muscu
C 4	61.4	95.9	141790	10 AL671759	AL671759 Mouse DNA
5	45.6	71.2	1477	9 AF070546	AF070546 Homo sapi
C 6	45.6	71.2	2710	11 BV178942	BV178942 sqmm10325
7	45.6	71.2	2944	9 BC065263	BC065263 Homo sapi
8	45.6	71.2	3049	9 AK096161	AK096161 Homo sapi
9	45.6	71.2	3495	9 BC029686	BC029686 Homo sapi
10	45.6	71.2	3554	9 BC046565	BC046565 Homo sapi
11	45.6	71.2	4809	9 BC080578	BC080578 Homo sapi
12	45.6	71.2	4880	9 HSM803946	HSM803946 Homo sapi
13	45.6	71.2	5023	9 HSM803947	HSM803947 Homo sapi
14	45.6	71.2	22255	6 AR409342	AR409342 Sequence
15	45.6	71.2	22255	6 AR239608	AR239608 Sequence
16	45.6	71.2	115756	9 HSDJ622L5	HSDJ622L5 Sequence
17	41.4	64.7	4722	6 AR139533	AR139533 Human DNA
18	41.4	64.7	4722	6 AR374685	AR374685 Sequence
19	41.4	64.7	4722	6 AR409320	AR409320 Sequence

20	41.4	64.7	4722	6 AX239571	AX239571 Sequence
21	41.4	64.7	4722	6 BD056447	BD056447 Novel low
22	39.2	61.3	4697	6 AR199536	AR199536 Sequence
23	39.2	61.3	4697	6 AR374688	AR374688 Sequence
24	39.2	61.3	4697	6 AR409323	AR409323 Sequence
25	39.2	61.3	4697	6 AX239574	AX239574 Sequence
26	39.2	61.3	4697	6 BD056450	BD056450 Novel low
27	31	48.4	204697	2 AC120995	AC120995 Rattus no
C 28	31	48.4	236070	2 AC112571	AC112571 Rattus no
C 29	30	46.9	155580	2 AC118845	AC118845 Rattus no
30	29.6	46.2	229213	10 AL772138	AL772138 Mouse DNA
31	29.2	45.6	218924	10 AC124567	AC124567 Mus muscu
C 32	29	45.3	133265	10 AL954355	AL954355 Mouse DNA
C 33	29	45.3	142667	10 AC118154	AC118154 Rattus no
C 34	29	45.3	216851	2 AC118154	AC118154 Rattus no
35	29	45.3	220407	2 AC121626	AC121626 Rattus no
36	29	45.3	233944	2 AC106583	AC106583 Rattus no
C 37	28.8	45.0	3742	5 BC077481	BC077481 Xenopus l
C 38	28.6	44.7	170788	2 AC139387	AC139387 Rattus no
C 39	28.6	44.7	244536	2 AC099176	AC099176 Rattus no
C 40	28.6	44.7	334320	2 AC113256	AC113256 Rattus no
C 41	28.4	44.4	378	6 C0656202	C0656202 Sequence
42	28.4	44.4	182389	9 AP002829	AP002829 Homo sapi
43	28.4	44.4	183090	2 AC121175	AC121175 Rattus no
C 44	28.4	44.4	196832	9 AC018694	AC018694 Homo sapi
C 45	28.4	44.4	208326	2 AC022888	AC022888 Homo sapi

ALIGNMENTS

AC111774 245095 bp DNA linear HTG 09-NOV-2002
Rattus norvegicus clone CH230-171G2, WORKING DRAFT SEQUENCE, 10
unordered pieces.

AC111774.4 GI:24818709
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 245095)

AUTHORS

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falla, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hammon, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpethy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowig, C., Kraft, C. I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lohenshuwa, L., Loulseghe, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Naik, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarinmoye, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Streimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 245095)
Worley, K.C.

Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245095)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 9, 2002 this sequence version replaced gi:23321701.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GOBQ
Center clone name: CH230-171G2
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 153773 bases at least Q40
Consensus quality: 160170 bases at least Q30
Consensus quality: 164659 bases at least Q20
Estimated insert size: 160182; sum-of-contigs estimation
Quality coverage: 10x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 10 'contigs'. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will

* be preserved.
* 1 69762: contig of 69762 bp in length
* 69862: gap of unknown length
* 69863 69862: contig of 3827 bp in length
* 69863 73789: gap of unknown length
* 73790 88447: contig of 14658 bp in length
* 88448 88447: gap of unknown length
* 88448 124613: contig of 36066 bp in length
* 124614 124713: gap of unknown length
* 124714 179270: contig of 54557 bp in length
* 179271 179370: gap of unknown length
* 179371 224496: contig of 45126 bp in length
* 224497 224597: gap of unknown length
* 224597 228137: contig of 3540 bp in length
* 228137 228237: gap of unknown length
* 228237 239465: contig of 11229 bp in length
* 239466 243842: gap of unknown length
* 243843 243943: contig of 4277 bp in length
* 243943 245095: gap of unknown length
* 245095 245095: contig of 1153 bp in length.

FEATURES
source
1. 245095
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-171G2"
/note="wgs_contig"
88548..89924
/note="wgs_contig"
91432..93819
/note="wgs_contig"
116848..118080
/note="wgs_contig"
124714..126005
/note="wgs_contig"
228237..229783
/note="wgs_contig"
239566..240945
/note="wgs_contig"

ORIGIN
Query Match 100.0%; Score 64; DB 2; Length 245095;
Best Local Similarity 100.0%; Pred. No. 6.3e-11;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATGACCTCATTATAGGACCAAGACCTGTTGGTTCTTAGATTGTTAGCTTTTCTC 60
Db 148400 TCATGACCTCATTATAGGACCAAGACCTGTTGGTTCTTAGATTGTTAGCTTTTCTC 148459

QY 61 TAGA 64
Db 148460 TAGA 148463

RESULT 2
BC060227
LOCUS
DEFINITION
BC060227.1 GI:38614270
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4702)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Dlatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:20:29 ; Search time 266 Seconds
(without alignments)

1424.299 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64

Sequence: 1 tcatgacctcatttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.6	71.2	1926	6	Abk35029 Human cDN
2	45.6	71.2	2145	13	Adp54990 Human PRO
3	45.6	71.2	5095	8	Acc46153 Human dit
4	45.6	71.2	6577	8	Acc46130 Human dit
5	45.6	71.2	22255	5	Aah26497 Human low
6	45.6	71.2	115756	8	Adi33448 Human DNA
7	41.4	64.7	4722	2	Aav32836 Rabbit lo
8	41.4	64.7	4722	5	Aah26491 Rabbit lo
9	39.2	61.3	4697	2	Aav32839 Human low
10	39.2	61.3	4697	5	Aah26496 Human low
11	38	59.4	635	13	Adq52076 Novel can
12	33.6	52.5	268	2	Aat33888 Human gen
13	27.2	42.5	1024	6	Abx66285 Helicobac
14	27.2	42.5	1024	6	Abx65642 Helicobac
15	27.2	42.5	1024	6	Abx66280 Helicobac
16	27.2	42.5	1035	2	Aax14487 H. pylori
17	27.2	42.5	6325	8	Abz10086 Haematopo
18	26.4	41.2	569	11	Acn87401 Breast ca
19	26.4	41.2	2418	6	Abq70768 Listeria
20	26.4	41.2	2875	12	Adm91308 DNA homol

C 21.	26.4	41.2	25800	12	ADQ97464	Adq97464 Mouse can
C 22	26.4	41.2	110000	6	ABA03041_21	Continuation (22 o
C 23	26.2	40.9	2730	2	AAT32583	Aat32583 Alfalfa i
C 24	26.2	40.9	15690	6	ABK15038	Abk15038 Canine di
C 25	26.2	40.9	15890	6	ABK15000	Abk15000 Canine di
C 26	26.2	40.9	18826	6	ABK15039	Abk15039 Canine di
C 27	25.8	40.3	16825	4	ABK42608	Abk42608 Genomic s
C 28	25.8	40.3	16825	9	ADB60764	ADB60764 Connectiv
C 29	25.8	40.3	23243	9	ADA02567	Ada02567 Mouse Gat
C 30	25.8	40.3	23243	10	ADB72305	ADB72305 Mouse Gat
C 31	25.8	40.3	23243	10	ADB95815	ADB95815 Mouse Gat
C 32	25.8	40.3	26997	4	AAS46748	Aas46748 Tumour su
C 33	25.8	40.3	192427	10	ADL13825	Adl13825 Osteoarth
C 34	25.8	40.3	310268	13	ABD32548	ABD32548 Human can
C 35	25.6	40.0	1754	10	ADB63505	Abd63505 Human cDN
C 36	25.6	40.0	6325	8	ABZ10232	Abz10232 Haematopo
C 37	25.6	40.0	110000	4	AAK95240_05	Continuation (6 of
C 38	25.6	40.0	110000	4	AAK96733_05	Continuation (6 of
C 39	25.6	40.0	110000	6	ABT00010_05	Continuation (6 of
C 40	25.6	40.0	110000	6	ABT01503_05	Continuation (6 of
C 41	25.6	40.0	110000	12	ADH77485_05	Continuation (6 of
C 42	25.4	39.7	28449	12	ADQ97240	Adq97240 Mouse can
C 43	25.4	39.7	256294	13	ABD33020	Abd33020 Mouse can
C 44	25.2	39.4	275	4	AAK83205	Aak83205 Human imm
C 45	25.2	39.4	427	6	ABL38563	Ab138563 Human col

ALIGNMENTS

RESULT 1

ABK35029

ID ABK35029 standard; cDNA; 1926 BP.

XX ABK35029;

AC ABK35029;

DT 08-MAY-2002 (first entry)

XX Human cDNA encoding secreted protein #167.

DE Human cDNA encoding secreted protein #167.

XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;

KW viral infection; bacterial infection; fungal infection; diabetes; asthma;

KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;

KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;

KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;

KW tissue regeneration; wound healing; burn; haematopoiesis;

KW myeloid cell deficiency; lymphoid cell deficiency.

XX Homo sapiens.

OS Homo sapiens.

XX WO200177288-A2.

PN 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US010224.

XX 06-APR-2000; 2000US-0195582P.

XX (GEMY) GENETICS INST INC.

XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;

XX Gulukota K, Graham JR;

XX WPI; 2002-179321/23.

XX Five hundred and ninety two polynucleotides derived from a variety of

XX human tissue sources which encode secreted proteins, useful for treating

XX immune deficiencies and disorders such as autoimmune disorders.

XX Claim 1; Page 152; 372pp; English.

XX The invention relates to 592 polynucleotides which have been derived from

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:25:03 ; Search time 97 Seconds
(without alignments)
1079.605 Million cell updates/sec

Title: US-10-663-418-71
Perfect score: 64
Sequence: 1 tcatgacctatttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgm2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.6	71.2	22255	4	US-09-616-289-51
2	41.4	64.7	4722	3	US-08-979-608A-14
3	41.4	64.7	4722	4	US-03-517-849-14
4	41.4	64.7	4722	3	US-09-616-289-14
5	39.2	61.3	4697	3	US-08-979-608A-17
6	39.2	61.3	4697	4	US-09-517-849-17
7	39.2	61.3	4697	4	US-09-616-289-17
8	28.8	45.0	601	4	US-09-949-016-171883
9	27.6	43.1	601	4	US-09-949-016-171884
10	27.6	43.1	232547	4	US-09-949-016-16603
11	26.4	41.2	601	4	US-09-949-016-149477
12	26.4	41.2	156942	4	US-09-949-016-12227
13	26.4	41.2	156950	4	US-09-949-016-15946
14	26.4	41.2	300598	4	US-09-949-016-11868
15	26.4	41.2	302604	4	US-09-949-016-14589
16	26.4	41.2	302604	4	US-09-949-016-14589
17	26.4	41.2	308362	4	US-09-949-016-17119
18	26.2	40.9	2730	1	US-08-339-129-1
19	25.8	40.3	601	4	US-09-949-016-158019
20	25.8	40.3	601	4	US-09-949-016-158020
21	25.8	40.3	36620	4	US-09-949-016-16150
22	25.8	40.3	87352	4	US-09-949-016-12053
23	25.8	40.3	87352	4	US-09-949-016-12721
24	25.8	40.3	87352	4	US-09-949-016-15692
25	25.8	40.3	87352	4	US-09-949-016-15693
26	25.4	39.7	601	4	US-09-949-016-124839
27	25.4	39.7	601	4	US-09-949-016-124840

```

28 25.4 39.7 80411 4 US-09-949-016-15777 Sequence 15777, A
29 25.4 39.7 192302 4 US-09-949-016-15270 Sequence 15270, A
c 30 25.2 39.4 254964 4 US-09-949-016-12583 Sequence 12583, A
c 31 25.2 39.4 254964 4 US-09-949-016-17392 Sequence 17392, A
32 25.2 39.4 392000 4 US-10-027-983-11 Sequence 11, Appl
33 25 39.1 601 4 US-09-949-016-130194 Sequence 130194, A
34 25 39.1 212449 4 US-09-949-016-15419 Sequence 15419, A
c 35 24.8 38.8 601 4 US-09-949-016-24772 Sequence 24772, A
c 36 24.8 38.8 601 4 US-09-949-016-24773 Sequence 24773, A
c 37 24.8 38.8 601 4 US-09-949-016-24774 Sequence 24774, A
c 38 24.8 38.8 601 4 US-09-949-016-35142 Sequence 35142, A
c 39 24.8 38.8 601 4 US-09-949-016-35143 Sequence 35143, A
c 40 24.8 38.8 601 4 US-09-949-016-35144 Sequence 35144, A
c 41 24.8 38.8 601 4 US-09-949-016-35171 Sequence 35171, A
c 42 24.8 38.8 601 4 US-09-949-016-35172 Sequence 35172, A
c 43 24.8 38.8 601 4 US-09-949-016-35173 Sequence 35173, A
c 44 24.8 38.8 601 4 US-09-949-016-173472 Sequence 173472, A
c 45 24.8 38.8 601 4 US-09-949-016-173473 Sequence 173473, A

```

ALIGNMENTS

```

RESULT 1
US-09-616-289-51
; Sequence 51, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 22255
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-616-289-51

```

Query Match 71.2%; Score 45.6; DB 4; Length 22255;
Best Local Similarity 92.3%; Pred. No. 7.6e-06;
Matches 48; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 12 TTTTAGGACCAAGAGCTGCTGCTTCTTAGATTGTTAGCTTTTCTCTAG 63
Db 19382 TTTTAGGACCAAGATCTGCTTCTTAGATTGTTAGCTTTTCTCTAG 19433

```

```

RESULT 2
US-08-979-608A-14
; Sequence 14, Application US/08979608A
; Patent No. 6355451
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:58:43 ; Search time 300 Seconds
(without alignments)
1295.158 Million cell updates/sec

Title: US-10-663-418-71
Perfect score: 64
Sequence: 1 tcagacacatttaggac.....ttgttagctttttcttaga 64

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5633728 seqs, 3035525691 residues

Total number of hits satisfying chosen parameters: 11267456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	45.6	71.2	1926	US-09-822-849A-167
2	45.6	71.2	22255	US-09-976-740-51
3	45.6	71.2	22255	US-10-023-529-51
4	45.6	71.2	22255	US-10-023-523-51
5	45.6	71.2	22255	US-10-616-187-51
6	45.6	71.2	22255	US-10-671-242-51
7	41.4	64.7	4722	US-09-962-055-14
8	41.4	64.7	4722	US-09-976-740-14
9	41.4	64.7	4722	US-10-023-529-14
10	41.4	64.7	4722	US-10-023-523-14
11	41.4	64.7	4722	US-10-616-187-14

12	41.4	64.7	4722	17	US-10-671-242-14	Sequence 14, Appl
13	39.2	61.3	4697	9	US-09-962-055-17	Sequence 17, Appl
14	39.2	61.3	4697	9	US-09-976-740-17	Sequence 17, Appl
15	39.2	61.3	4697	13	US-10-023-529-17	Sequence 17, Appl
16	39.2	61.3	4697	13	US-10-023-523-17	Sequence 17, Appl
17	39.2	61.3	4697	17	US-10-616-187-17	Sequence 17, Appl
18	39.2	61.3	4697	17	US-10-671-242-17	Sequence 17, Appl
c 19	28.4	44.4	378	17	US-10-242-535A-1128	Sequence 1128, Ap
c 20	28.4	44.4	378	17	US-10-085-783A-1128	Sequence 1128, Ap
c 21	27.2	42.5	1035	10	US-09-882-227-401	Sequence 401, App
c 22	27.2	42.5	6325	18	US-10-473-126-226	Sequence 226, App
c 23	27.2	42.5	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 24	26.8	41.9	818	13	US-10-027-632-152420	Sequence 152420,
c 25	26.8	41.9	818	17	US-10-027-632-152420	Sequence 152420,
c 26	26.4	41.2	569	14	US-10-198-846-8551	Sequence 8551, Ap
c 27	26.4	41.2	739	13	US-10-027-632-171545	Sequence 171545,
c 28	26.4	41.2	739	17	US-10-027-632-171545	Sequence 171545,
c 29	26.4	41.2	2418	17	US-10-398-221-3581	Sequence 3581, Ap
c 30	26.2	40.9	673	13	US-10-027-632-7733	Sequence 7733, Ap
c 31	26.2	40.9	673	13	US-10-027-632-7734	Sequence 7734, Ap
c 32	26.2	40.9	673	13	US-10-027-632-7735	Sequence 7735, Ap
c 33	26.2	40.9	673	17	US-10-027-632-7735	Sequence 7735, Ap
c 34	26.2	40.9	673	17	US-10-027-632-7734	Sequence 7734, Ap
c 35	26.2	40.9	673	17	US-10-027-632-7735	Sequence 7735, Ap
c 36	26	40.6	492	13	US-10-027-632-267996	Sequence 267996,
c 37	26	40.6	492	13	US-10-027-632-267997	Sequence 267997,
c 38	26	40.6	492	17	US-10-027-632-267996	Sequence 267996,
c 39	26	40.6	492	17	US-10-027-632-267997	Sequence 267997,
40	25.8	40.3	16825	9	US-09-764-847-1495	Sequence 1495, Ap
41	25.8	40.3	16825	14	US-10-092-154-1495	Sequence 1495, Ap
42	25.8	40.3	23243	17	US-10-052-482-73	Sequence 73, Appl
43	25.8	40.3	26997	17	US-10-221-714A-474	Sequence 474, App
44	25.8	40.3	301692	17	US-10-428-487-11	Sequence 11, Appl
45	25.8	40.3	310268	18	US-10-367-094-195	Sequence 195, App

ALIGNMENTS

RESULT 1

US-09-822-849A-167
; Sequence 167, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-167

Query Match 71.2%; Score 45.6; DB 9; Length 1926;
Best Local Similarity 92.3%; Pred. No. 1e-05; 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 4; Indels 0;

QY 12 TTTTAGACCAAGAGCTGTGTTCTTCTAGATTGTAGCTTTTCTCTAG 63
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 24, 2005, 07:21:49 ; Search time 1820 Seconds
(without alignments)
1338.524 Million cell updates/sec

Title: US-10-663-418-71

Perfect score: 64

Sequence: 1 tcacgacctatttaggac.....ttgttagctttttctctaga 64

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68473088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	64	100.0	566	1	AA925303	UI-R-A1-e
C 2	64	100.0	649	5	BQ192706	UI-R-DRI-
C 3	64	100.0	659	6	CA503893	UI-R-F00-
C 4	64	100.0	683	6	CB322784	UI-R-DYO-
5	64	100.0	802	7	CO395194	AGENCOURT
6	63	98.4	419	2	BF555411	UI-R-A1-e
7	63	98.4	477	1	AA869795	qqlsh08.r
8	63	98.4	546	5	BX516797	EX516797
9	63	98.4	641	6	CA880144	K0980R08-
10	63	98.4	641	6	CB058161	NISC jx01
11	63	98.4	1009	6	CB208950	AGENCOURT
12	63	98.4	1167	2	BF180039	601806391
13	61.4	95.9	519	1	AA170252	ms87e12.r
14	61.4	95.9	540	1	AA915373	vz29d01.r
15	61.4	95.9	647	6	CA450819	UI-M-FV0-
C 16	61.4	95.9	651	5	BP762179	BP762179
17	61.4	95.9	651	7	CF727202	UI-M-H80-
C 18	61.4	95.9	652	5	BQ176939	UI-M-DJ2-
C 19	61.4	95.9	655	5	BP761959	BP761959
C 20	61.4	95.9	675	7	CO042840	UI-M-GH0-
21	61.4	95.9	694	2	BF120794	601758189
22	61.4	95.9	698	3	AK011182	Mus muscu
23	61.4	95.9	712	6	BY710509	BY710509
24	61.4	95.9	717	6	CA324014	UI-M-FY0-

C 25	61.4	95.9	748	7	CO042637	UI-M-GH0-
C 26	61.4	95.9	760	7	CO042337	UI-M-GH0-
27	61.4	95.9	856	2	BB038251	BB038251
28	59.8	93.4	454	6	BY597346	BY597346
29	59.8	93.4	543	5	BU609826	UI-M-DJ2-
30	59.8	93.4	755	7	CN529567	UI-M-HQ0-
31	54.4	85.0	174	1	AI608525	vc73b12.y
32	53.4	83.4	650	4	EG867325	602788676
33	47.2	73.8	309	2	AW487474	AW487474
34	47.2	73.8	340	2	AW446148	AW446148
35	47.2	73.8	481	2	BE755989	BE755989
36	47.2	73.8	482	4	BM090178	BM090178
37	47.2	73.8	513	4	BI774054	BI774054
C 38	47.2	73.8	640	7	CK953122	CK953122
C 39	47.2	73.8	667	7	CK950080	CK950080
40	47.2	73.8	692	7	CK953458	CK953458
41	46.2	72.2	548	1	AV604019	AV604019
42	45.6	71.2	364	7	T99127	T99127
43	45.6	71.2	432	4	BM766498	BM766498
44	45.6	71.2	489	4	BM689900	BM689900
45	45.6	71.2	505	2	BE185069	BE185069

ALIGNMENTS

RESULT 1
AA925303/c

LOCUS

DEFINITION

UI-R-A1-ek-e-09-0-UI-s1

UI-R-A1-ek-e-09-0-UI-3'

similar to GI1768059[gb|AA184410|AA184410

mt34f05.r1 Soares mouse 3NDMS Mus musculus cDNA clone 622977 5',

mRNA sequence.

ACCESSION

AA925303

VERSION

AA925303.1

KEYWORDS

EST

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 566)

Bonaldi,M.F., Lemmon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

On Apr 21, 1998 this sequence version replaced gi:3072439.

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road . 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence tag present in the cDNA between the NotI site and the

oligo-dT track served to identify it as a clone from the normalized

adult muscle library. cDNA Library Preparation: M. Fatima Bonaldi,

Ph.D. Clone distribution: clones will be available through Research

Genetics This clone is also available through the I.M.A.G.E.

Consortium at LLNL (info:image.llnl.gov). IMAGE ID=1771492

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1..566

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-A1-ek-e-09-0-UI"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

AA925303 566 bp mRNA linear EST 04-JUL-1999
UI-R-A1-ek-e-09-0-UI-s1 UI-R-A1 Rattus norvegicus cDNA clone
UI-R-A1-ek-e-09-0-UI-3' similar to GI1768059[gb|AA184410|AA184410
mt34f05.r1 Soares mouse 3NDMS Mus musculus cDNA clone 622977 5',
mRNA sequence.

ACCESSION

AA925303

VERSION

AA925303.1

KEYWORDS

EST

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 566)

Bonaldi,M.F., Lemmon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

On Apr 21, 1998 this sequence version replaced gi:3072439.

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road . 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence tag present in the cDNA between the NotI site and the

oligo-dT track served to identify it as a clone from the normalized

adult muscle library. cDNA Library Preparation: M. Fatima Bonaldi,

Ph.D. Clone distribution: clones will be available through Research

Genetics This clone is also available through the I.M.A.G.E.

Consortium at LLNL (info:image.llnl.gov). IMAGE ID=1771492

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1..566

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-A1-ek-e-09-0-UI"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

Seq primer: M13 Forward
POLYA=yes.

FEATURES
source

Location/Qualifiers
1..649
/organism="Rattus norvegicus"
/mol_type="mrna"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-DRI-cla-d-11-0-UI"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="UI-R-DRI"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-DRI library is a normalized Rat Osteoblast library (nREO) constructed in pT737 vector according to the procedure described by Bonaldo, Lennon & Soares (Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery, Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag AAGATATCAA between the Not I cloning site and d18 stretch. The Rat Osteoblast tissue was provided by Lian & Stein of the University of Massachusetts Medical School.
TAG_TISSUE=osteoblast
TAG_LIB=UI-R-DRI
TAG_SEQ=AAGATATCAA"

ORIGIN

Query Match 100.0%; Score 64; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 60
|||||
565 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 506

QY 61 TAGA 64
|||||
505 TAGA 502

/clone lib="UI-R-A1"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-A1 library is a substracted library derived from the UI-R-A0 library. The UI-R-A0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, and muscle. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The substracted library (UI-R-A1) was constructed as follows: PCR amplified cDNA inserts from a pool of approximately 3,840 UI-R-A0 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-A0 library in the form of single-stranded circles. The remaining single-stranded circles (substracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-A1 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)."

ORIGIN

Query Match 100.0%; Score 64; DB 1; Length 566;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 60
|||||
565 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 506

QY 61 TAGA 64
|||||
505 TAGA 502

ORIGIN

Query Match 100.0%; Score 64; DB 5; Length 649;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 60
|||||
575 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATTGTTAGCTTTTCTC 516

QY 61 TAGA 64
|||||
515 TAGA 512

RESULT 3
CA503893/c
LOCUS
DEFINITION
UI-R-FJ0-cpv-o-09-0-UI.s1 UI-R-FJ0 Rattus norvegicus cDNA clone
UI-R-FJ0-cpv-o-09-0-UI 3', mRNA sequence.

ACCESSION
CA503893
VERSION
CA503893.1 GI:24994847
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 659)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

RESULT 2
BQ192706/c
LOCUS
DEFINITION
UI-R-DRI-cla-d-11-0-UI.s1 UI-R-DRI Rattus norvegicus cDNA clone
UI-R-DRI-cla-d-11-0-UI 3', mRNA sequence.

ACCESSION
BQ192706
VERSION
BQ192706.1 GI:20368257
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 (bases 1 to 649)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
normalized osteoblast library cDNA Library Preparation: M.B. Soares
Lab Clone distribution: clones will be available through Research
Genetics (www.regen.com)

Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1. 659

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-FJ0-cpv-o-09-0-UI"
/tissue_type="embryo"
/dev_stage="embryo"
/lab_host="DHI0B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-R-FJ0"
/notes="Vector: pXX-Asc; Site 1: EcoR I; Site 2: Not I; UI-R-FJ0 is a cDNA library containing the following tissue(s): rat embryo. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)18 tail. The sequence tag for this library is CATCTCTACT. This library was created for the University of Iowa Program for Rat Gene Discovery and Mapping (Val Sheffield, Bento Soares and Tom Casavant)
TAG_TISSUE=rat-embryo
TAG_LIB=UI-R-FJ0
TAG_SEQ=CATCTCTACT"

ORIGIN

Query Match 100.0%; Score 64; DB 6; Length 659;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTTGTTCTTAGATTGTAGCTTTTCTC 60
|||||
DB 569 TCATGACCTCATTTAGGACCAAGAGCTGTGTTGTTCTTAGATTGTAGCTTTTCTC 510
QY 61 TAGA 64
|||||
DB 509 TAGA 506

RESULT 4

CB322784/c
LOCUS
DEFINITION
UI-R-DY0-circ-d-14-0-UI.s1 NCI CGAP_DY0 Rattus norvegicus cDNA clone
IMAGE:7329856 3', mRNA sequence.
CB322784 683 bp mRNA linear EST 11-AUG-2004
CB322784
IMAGE:7329856 3', mRNA sequence.
CB322784.1 GI:28857442
EST.
Rattus norvegicus (Norway rat)
ORGANISM

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through IMAGE (http://image.llnl.gov)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1. 683

/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="IMAGE:7329856"
/dev_stage="adult"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="NCI CGAP DY0"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; UI-R-DY0 is a non-normalized rat cartilage library (RC) constructed in pT73 PAC vector according to the procedure described by Bonaldo, Lennon & Soares (Genome Research 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag CTAATGGAGC between the Not I cloning site and dt18 stretch. The Rat cartilage tissue was provided by Dr Jeff Stevens at the University of Iowa.
TAG_TISSUE=cartilage
TAG_LIB=UI-R-DY0
TAG_SEQ=CTAATGGAGC"

ORIGIN

Query Match 100.0%; Score 64; DB 6; Length 683;
Best Local Similarity 100.0%; Pred. No. 2.1e-10;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTTGTTCTTAGATTGTAGCTTTTCTC 60
|||||
DB 570 TCATGACCTCATTTAGGACCAAGAGCTGTGTTGTTCTTAGATTGTAGCTTTTCTC 511
QY 61 TAGA 64
|||||
DB 510 TAGA 507

RESULT 5

LOCUS
DEFINITION
CO395194 802 bp mRNA linear EST 01-JUL-2004
AGENCOURT 27849307 NIH_MGC_252 Rattus norvegicus cDNA clone
IMAGE:7310479 5', mRNA sequence.
CO395194
CO395194.1 GI:49577110
EST.
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus;
1 (bases 1 to 802)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
College of Wisconsin
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
College of Wisconsin
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
 Plate: L1AM15356 row: m column: 05
 High quality sequence stop: 654.
 Location/Qualifiers
 1. .802
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAG:7310479"
 /tissue types="Ovary, strain - Norway Line3. Age8 weeks.
 Tissue was snap-frozen adn transferred in -70C. RNase free
 for the entire procedure"
 /lab_hosts="DH10B Toran"
 /clone_lib="NIH_MGC 252"
 /notes="Organ: Ovary; Vector: pDONR 201; Site 1: NotI;
 Site 2: NotI; RNA obtained from female ovaries animals at
 8 wk old. Tissues were snap-frozen and kept at -80C for
 two days before RNA extraction and purification
 (TRI-reagent method). cDNA was primed using oligo-dT
 primer: 5'-pGACTAGTCTAGATCGGCGGCCGCT(7)25-3' and
 cloned into the EcoRV/NotI sites of pExpress-1.
 Size-selection 1.25kb resulted in an average insert size
 of 1.7kb. This primary library is not normalized
 (normalized library is NIH_MGC 252) and was constructed by
 Express Genomics (Frederick, MD). Note: this is a NIH_MGC
 library"

ORIGIN
 Query Match 100.0%; Score 64; DB 7; Length 802;
 Best Local Similarity 100.0%; Pred. No. 2.2e-10;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCATGACCTCATTTAGACCAAGAGCTGTGTGGTTCTTAGATGTTAGCTTTTC 60
 Db 536 TCATGACCTCATTTAGACCAAGAGCTGTGTGGTTCTTAGATGTTAGCTTTTC 595
 QY 61 TAGA 64
 Db 596 TAGA 599

RESULT 6
 BF555411
 LOCUS
 DEFINITION
 UI-R-A1-ek-e-09-0-UI-r1 UI-R-A1 Rattus norvegicus cDNA clone
 UI-R-A1-ek-e-09-0-UI 5', mRNA sequence.
 ACCESSION
 BF555411
 VERSION
 BF555411.1 GI:11665141
 KEYWORDS
 EST.
 ORGANISM
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 419)
 REFERENCE
 AUTHORS
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL
 MEDLINE
 97044477
 PUBMED
 8889548
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.reagen.com)
 This clone is also available through the I.M.A.G.E. Consortium at
 LLNL (info@image.llnl.gov). IMAGE ID= 1771492
 Seq primer: M13 Forward.

Location/Qualifiers
 1. .419
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strains="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-A1-ek-e-09-0-UI"
 /dev_stages="adult"
 /lab_hosts="DH10B (Life Technologies)"
 /clone_lib="UI-R-A1"
 /notes="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-A1
 library is a subtracted library derived from the UI-R-A0
 library. The UI-R-A0 library consisted of a mixture of
 individually tagged normalized libraries constructed from
 rat placenta, adult lung, brain, liver, kidney, heart,
 spleen, ovary, and muscle. The tag is a string of 3-5
 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (UI-R-A1) was constructed as follows: PCR
 amplified cDNA inserts from a pool of approximately 3,840
 UI-R-A0 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-A0
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-A1 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

ORIGIN
 Query Match 98.4%; Score 63; DB 2; Length 419;
 Best Local Similarity 98.4%; Pred. No. 4.2e-10;
 Matches 63; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATGTTAGCTTTTC 60
 Db 308 TCATGACCTCATTTAGGACCAAGAGCTGTGTGGTTCTTAGATGTTAGCTTTTC 367
 QY 61 TAGA 64
 Db 368 TAGA 371

RESULT 7
 AA869795
 LOCUS
 DEFINITION
 vq15h08.r1 Barstead stromal cell line MPLRB8 Mus musculus CDNA
 clone IMAGE:1094367 5', mRNA sequence.
 ACCESSION
 AA869795
 VERSION
 AA869795.1 GI:2965240
 KEYWORDS
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 477)
 REFERENCE
 AUTHORS
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE
 The WashU-HMI Mouse EST Project
 JOURNAL
 UNPUBLISHED (1996)
 COMMENT
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810